

SUPPLEMENTARY MATERIAL

corresponding to:

**Differential expression of *arid5b* isoforms
in *Xenopus laevis* pronephros**

RONAN LE BOUFFANT, ANNE-CLAIRE CUNIN, ISABELLE BUISSON, JÉRÔME CARTRY,

JEAN-FRANÇOIS RIOU and MURIEL UMBHAUER*

*Address correspondence to: Muriel Umbhauer. Laboratoire de Biologie du Développement UMR7622, CNRS, UPMC, 9 quai Saint-Bernard, case 24, 75005, Paris France. Tel: (33)-1-4427-3918: fax: (33)-1-4427-3445. e-mail: muriel.umbhauer@upmc.fr

1 10 20 30 40 50 60
X.laevis MEPNSLK WVGSSCGLHGPIFYKAFQFHLEN RARILSLGDFFLVRC KHEPVCVAELQLLWEERTSR
X.tropicalis MEPNSLK WVGSSACGLHGPIFYKAFQFHLEN RARILSLGDFFLVRC KADPEVCIAELQLLWEERTSR
G.gallus MERSALQ WVGAPCGSHGPIVYFRFRFQRRGG RARVLSLGDFFVRC RAEPEACIAELQLLWEERTSR
H.sapiens MEPNSLQ WVGSPCGLHGPIFYKAFQFHLEG KPRILSLGDFFFVRC TPKDPICIAELQLLWEERTSR
M.musculus MEPNSLQ WVGSPCGLHGPIFYKAFQFHLEG KPRILSLGDFFFVRC TPKDPICIAELQLLWEERTSR
D.terio
consensus>50 mepnslq.wvgspcglhgpyifykafqfhleg..rarilslgdfffvrc.p.epvciaelqllweertsr

70 80 90 100 110 120 130
X.laevis QLLSSSKLYFLPEDTFPKGRNSHGEHEVIIVSEKVIIVRLEDLVKWAHSDFSKWNYGLKALPVE...LKEKEL
X.tropicalis QLLSSSKLYFLPEDTFPKGRNSHGEHEVIIVSEKVIIVRLEDLVKWAHSDFSKWNYGLKALPVE...LKEKEL
G.gallus QLLSSAKLYFLPEDTPQGRSDHGEDEVIIVSEKVIIVKLEDLAKWAOSDFSKWNYGLKALPVE...PMDV
H.sapiens QLLSSSKLYFLPEDTPQGRSDHGEDEVIIVSEKVIIVKLEDLVKWAHSDFSKWNYGLKALPVE...TEAL
M.musculus QLLSSSKLYFLPEDTPQGRSDHGEDEVIIVSEKVIIVKLEDLVKWAHSDFSKWNYGLKALPVE...TEAF
D.terio
consensus>50 qlsssklyflpedtpqgrnsdhgedeviavsekvivkledlvkwhsdfskw.cgl.a.pvk...lkel

140 150 160 170 180 190 200
X.laevis ARNGQRESLAKYRQSLVNSGLNFKDVLKEKALGEGEDKNVMVLSYPOYCRYRSILRLRQAEPSVLLAD
X.tropicalis ARNGQRESLAKYRQSLVNSGLNFKDVLKEKALGEGEDKNVMVLSYPOYCRYRSILRLRQAEPSVLLAD
G.gallus GRNGQRESLAKYRQSLVNSGLNFKDVLKEKALGEGEDKNVMVLSYPOYCRYRSILRLRQAEPSVLLAD
H.sapiens GRNGQRESLAKYRQSLVNSGLNFKDVLKEKALGEGEDKNVMVLSYPOYCRYRSILRLRQAEPSVLLAD
M.musculus GRNGQRESLAKYRQSLVNSGLNFKDVLKEKALGEGEDKNVMVLSYPOYCRYRSILRLRQAEPSVLLAD
D.terio ARNGQRESLAKYRQSLVNSGLNFKDVLKEKALGEGEDKNVMVLSYPOYCRYRSILRLRQAEPSVLLAD
consensus>50 gkngqkEaLlkyR#stLNSGLNFKD!lKEKA#LGEdee#knvm!LSYPQYCRYRSilkrIdqkPSSiLd

210 220 230 240 250 260 270
X.laevis QFVQALGGIIVAVSRNPFIIYCRDTHPTL!ENES!CDEFAPNLKGRPRKKK!LGI.PQRR#S.NGVKELAG
X.tropicalis QFVQALGGIIVAVSRNPFIIYCRDTHPTL!ENES!CDEFAPNLKGRPRKKK!LGI.PQRR#S.NGVKELAG
G.gallus QFVQALGGIIVAVSRNPFIIYCRDTHPTL!ENES!CDEFAPNLKGRPRKKK!LGI.PQRR#S.NGVKELAG
H.sapiens QFVQALGGIIVAVSRNPFIIYCRDTHPTL!ENES!CDEFAPNLKGRPRKKK!LGI.PQRR#S.NGVKELAG
M.musculus QFVQALGGIIVAVSRNPFIIYCRDTHPTL!ENES!CDEFAPNLKGRPRKKK!LGI.PQRR#S.NGVKELAG
D.terio HVVQALGGIIVAVSRNPFIIYCRDTHPTL!ENES!CDEFAPNLKGRPRKKK!LGI.PQRR#S.NGVKELAG
consensus>50 qfvqalGGIiavsrnnpqIIYCRDTHPTL!ENES!CDEFAPNLKGRPRKKK.c.pQRR#S.nGvkd.nn

280 290 300 310 320 330 340
X.laevis VCDGRFVAKVVKCEVKATLHKTKTPN.GNCKKILIEERKRAGEDDGGCKVDEGKABEQAFVLAALYKMKERR
X.tropicalis VCDGRFVAKVVKCEVKATLHKTKTPN.GNCKKILIEERKRAGEDDGGCKVDEGKABEQAFVLAALYKMKERR
G.gallus NSESAVAKVVKCEAKSALHKPKSNN.SNCKKGSSESRKAVG.....EGRABEQAFVLAALYKMKERR
H.sapiens NSDESAVAKVVKCEARSALHKPKNNH.NCKKSNSESRKAVG.....EGRABEQAFVLAALYKMKERR
M.musculus NCDGRFVAKVVKCEARSALHKPKNNH.NCKKSNSESRKAVG.....EGRABEQAFVLAALYKMKERR
D.terio GVEGRFVAKVVKCEVKATLHKTKTPN.GNCKKILIEERKRAGEDDGGCKVDEGKABEQAFVLAALYKMKERR
consensus>50 nc#gktvaKvk#aKsalpKpknnn.gnCKkis.E#kpk..ig.....#EcrA#EQAFVLAALYKMKERR

350 360 370 380 390 400 410
X.laevis TPIERIPYLGFKQINLWTFQAAQKLGGEYETITARRQWKHIYDELGGNPGSTSAATCRRHYERLILPYE
X.tropicalis TPIERIPYLGFKQINLWTFQAAQKLGGEYETITARRQWKHIYDELGGNPGSTSAATCRRHYERLILPYE
G.gallus TPIERIPYLGFKQINLWTFQAAQKLGGEYETITARRQWKHIYDELGGNPGSTSAATCRRHYERLILPYE
H.sapiens TPIERIPYLGFKQINLWTFQAAQKLGGEYETITARRQWKHIYDELGGNPGSTSAATCRRHYERLILPYE
M.musculus TPIERIPYLGFKQINLWTFQAAQKLGGEYETITARRQWKHIYDELGGNPGSTSAATCRRHYERLILPYE
D.terio TPIERIPYLGFKQINLWTFQAAQKLGGEYETITARRQWKHIYDELGGNPGSTSAATCRRHYERLILPYE
consensus>50 TPIERIPYLGFKQINLWTFQAAQKLGGEYETITARRQWKHIYDELGGNPGSTSAATCRRHYERLILPYE

420 430 440 450 460 470 480
X.laevis RFIKGEEDKPLPFAKPRKQENSSQEVVDMKVKICGTRKRIKNEPQKPKKED...AKGYDMAEVPPDEEDHLE
X.tropicalis RFIKGEEDKPLPFAKPRKQENSSQEVVDMKVKICGTRKRIKNEPQKPKKED...AKGYDMAEVPPDEEDHLE
G.gallus RFIKGEEDKPLPFAKPRKQENSSQEVVDMKVKISGTRKRIKNEPQKPKKED...AKGYDMAEVPPDEEDHLE
H.sapiens RFIKGEEDKPLPFAKPRKQENSSQEVVDMKVKISGTRKRIKNEPQKPKKED...AKGYDMAEVPPDEEDHLE
M.musculus RFIKGEEDKPLPFAKPRKQENSSQEVVDMKVKISGTRKRIKNEPQKPKKED...AKGYDMAEVPPDEEDHLE
D.terio RFIKGEEDKPLPFAKPRKQENSSQEVVDMKVKISGTRKRIKNEPQKPKKED...AKGYDMAEVPPDEEDHLE
consensus>50 RFIKGEEDKPLPFAKPRKQENSSQEVVDMKVKISGTRKRIKNEPQKPKKED...AKGYDMAEVPPDEEDHLE

490 500 510 520 530 540
X.laevis A..DEKNMFLDYDLEETKTSVDRSESVVSEV.NYPSPLENDELLEETVANKDHVTKDENSCQDPDPVDSLI
X.tropicalis A..DEKNMFLDYDLEETKTSVDRSESVVSEV.NYPSPLENDELLEETVANKDHVTKDENSCQDPDPVDSLI
G.gallus SA..DQKNMFLDYDLEETKTSVDRSESVVSEV.NYPSPLENDELLEETVANKDHVTKDENSCQDPDPVDSLI
H.sapiens TLISQKSLPEPLPAADMKKIEGYQFSAPLASRVDPKEDNETDGSNSSEKVAEEAGEKTPPLPSAP
M.musculus TL..NHKSAPEPLPAPEVKKPEGHKLLGARAPVSRADPEKANEEDDGSNSSEKVAEEAGEKTPPLPSAP
D.terio
consensus>50 ...dekn.pe..dae#tk..v#g.ed...e...sp...#nde.teq.sns#kv.e#...d..plpSlp

550 560 570 580 590 600 610
X.laevis HVKEINCRQT...DKQLQMPNEMTITTKRE.QTKEDYSDHLENDPEDVQLHVFPFAIQAP.QHDMLEEEK
X.tropicalis CVKEIDSRQT...DERLQMPNEMTITTKRE.QTKEDYSDHLENDPEDVQLHVFPFAIQAP.QHDMLEEEK
G.gallus MPPEEDTVDLDAVTVKRLHSSADTQEDTKPERRHKAPTESLENEPPEMFFTFVVP...TQSDMEEDDK
H.sapiens IAPKDSALVPGASKQPLTSPSALVDSKQESKLC.CFTESPESEPPQASFPSPFTTQPLANONETEDDDK
M.musculus IAPKDSALVPGASKQPLTSPSALVDSKQESKLC.CFTESPESEPPQASFPSPFTTQPLANONETEDDDK
D.terio
consensus>50 l.pe.ds..t...kql..s.etm.dtkqe.qic..%teslendped.pl..fpa.qpp...q.e.e#k

620 630 640 650 660 670 680

X.laevis IPDMPDYIANCTVVKVDPIGSDLKNPLDSNLLQNALKQNPKVYFVOTLDMLSDEKDTASMNDDSSFSYT
X.tropicalis IPDMPDYIANCTVVKMDPIGSSDLKNPLDSNLFQNALKQNPKVYLVO TLDMLN.EKEAFPSMNDDSSFSYT
G.gallus IPEMADYIANCTVVKVDQIGNEDI.....ANALKQNPKVLVVONFDMFK.EKELTGPSSMDDSTFGYT
H.sapiens IPEMADYIANCTVVKVDQIGSDDI.....ANALKQNPKVLVVONFDMFK.DKDLTGPMMENHGLNYT
M.musculus IPEMADYIANCTVVKVDQIGSDDI.....HTALKQNPKVLVVONFDMFK.DKDLTGPMMENHGLNYT
D.ierio VPSA.IANVEQSRPKKGNQV.....VMVLPFLOQKPVTSPEIIPERVPLKKEESCENFN
consensus>50 lPdmadyIANctvkvvdqIGsni.....hnalqktPkvlvvq.fdmfk.ekdl.gsnn##s.fn#t

690 700 710 720 730 740 750

X.laevis PLLYSRGNPGIMSP LAKKKLSQVSGASQPCNTPYGSPPLISKKKLSSKGEVSPSLQTHHSSNS..ES
X.tropicalis PLLYSRGNPGIMSP LAKKKLSQVSGASQPCNTPYGSPPLINKKKLSSKGEVSPSLQTHHSSNS..ES
G.gallus PLLYSRGNPGIMSP LAKKKLSQVSGAALSQSYGSPPLISKKKLNSRDELSSSLQGGHPVNS..DF
H.sapiens PLLYSRGNPGIMSP LAKKKLSQVSGASLSSSYGSPPLISKKKLIARDLCSLSQTHHGQST..DH
M.musculus PLLYSRGNPGIMSP LAKKKLSQVSGASLSSSYGSPPLISKKKLIAREDLCSLQGGHHSQSS..DH
D.ierio PLLYSRGNPGIMSP LAKKKLSQVSGTGLLNQVYGYGSPPLVSRRLSSSGTEVSSAGQSSSQVSSVETTS
consensus>50 PLLYsrGNPGIMSP LAKKKLSQVSGasls.nyPYGSPPL!skkklsrrd#vssslsqthhssns..ds

760 770 780 790 800 810 820

X.laevis AAINRPSVIOHVSQSFQKRSPEEKKTVDNDHYKNSMFGKVDSYCCDFARHQSVLADSVALKACVQECKRKM
X.tropicalis VAINRPSVIOHVSQSFQKRSPEEKKTVDNDHYKNSMFGKVDSYCCDFARHQSVLADSVALKASIQECKRKM
G.gallus IAINRPSVIOHVSQSFQKRS..BERKSINDVFKHDMLSKVDIQRCDFSQHLCSLAESEVYPTDIQCKRKM
H.sapiens MAVSRPSVIOHVSQSFQKRSPEERKTINDIFKHEKLSRSDHRCFSFSLHNP LADSVYLRKEIQEGKRM
M.musculus TAVSRPSVIOHVSQSFQKRSPEERKSINDIFKHKLSRSDAHRCGFSFSLHNP LADSVYLRKEIQEGKRM
D.ierio IVIRPSVIOHVSQSFQKRSPEERRSSTEGSOKDGCSEGEF.....VHSQTLITREPVKRRVDPHSSMPKM
consensus>50 ia!nRPSVIOHvQSFqkqk.sE#rksin#ifkhdmlskv#p.rcdfskH.ls.la#sYvLk.dieqck#Ks

830 840 850 860 870 880

X.laevis AKRAASNSNVPSFVAEFYSSPHLHLCRAAEHHLHNEBSAKFH SRDMF.RDLSEVSS..LHKH...
X.tropicalis AKRAASNSNVPSFLAEFYSSPHLHLCRAAEHHLHNEBSAKFH SRDMF.RDLSEVSS..LHKH...
G.gallus TKRALQHSNVPSFLAEFYSSPHLHLCRAAEHHLHNEBSAKFH SRDMF.RDLSEVSS..LHKH...
H.sapiens LKRALPHSHNVPSFLAEFYSSPHLHLCRAAEHHLHNEBSAKFH SRDMY.RDLSEVSS..LHKH...
M.musculus LKRALVSHANVPSFLAEFYSSPHLHLCRAAEHHLHNEBSAKFH SRDAY.QESPE.GAFLSHKKH...
D.ierio AEMPRFGQALPSFLSEFYSSPHLHLCRAAEHHLHNEBSAKFH SRDVTEDSEPAQCEPESQH...
consensus>50 aEkra.sns.vPsFla#FYSSPHLH.LyRqtEHLHne#sSk%.sR#m%.r#sEnvssf.shkh.ek...

890 900 910 920 930 940

X.laevis HYHASLHQHDKQ..NLHDDVDDOPTDLSLPKS.LHKLSTKIPGSSICHQPV.OODSKSHNPFOTPNKSTL
X.tropicalis HYHASLHQHDKQ..NLCDIVDDOPTDLSLPKS.LHKLATRIPGSSIAHHSV.OODAKSHNPFOTPNKSTL
G.gallus NYHPSLHQQEKK..ATVEASDDOPTDLSLPKS.LHKQTAKAPCPGLSHSSMAOQGGKSVISFOAASSQAV
H.sapiens NYLPSLHQDQKSAAAEAFTDDOPTDLSLPKN.PHKETGKVLG..LAHSTGPOEIKSISFOVLSGSO..
M.musculus NYLASLHQDQKVAAAEAFTDDOPTDLSLPKN.PHKLTSKVLG..LAHSTGSOEIKSASFOVNSO..
D.ierio NPSARLSQKQK..PPEERVTEOPTDLSLPKS.SPLKPLSTSTLGGIPHAATOODIKSIPFOAGVNSOSS
consensus>50 n%hasLhqq#Kk.a..ea.t##QPTDLSLPKS.phKlt.kipg..lahs...qq#.Kg.s.Fq..nsq..

950 960 970 980 990 1000 1010

X.laevis GLDCNPKACRVSPMTEPISKRHMDSIQRP.SK..TVKPDTLR..KVEGLVHPFSIGKTNTNHFEGAPRSLK
X.tropicalis GVDCTPKACRVSPMTEMSMKRNMDSIHRP.SK..TLKPDTVR..KVEGLVHPFSIGKTSTONFGPPRSLK
G.gallus GLDCNPKACRVSPMAMTAKKRHSELHRS.SKQAQRLNLR..KMEGMVHPITSRMSPONIGAAEPK
H.sapiens SRDCHPKACRVSPMMSGPKKYPELSRS.SKPHVRLNFR..KMEGMVHPILHRKMSPONIGAAEPK
M.musculus SRDCHPKACRVSPMMSGPKKYPELSRS.SKPHQRLNFR..KMEGMVHPILHRKMSPONIGAAEPK
D.ierio SVYHHPACRVSPMTEVSSKRYTESHKVLKLEKPNRGESEMGFKIDEMSRPLSKSSPONICARPLK
consensus>50 s.DchPkaCRVSPMtmS.pKk..#sl.rs.gk...vrl#n.r..Kv#gqvhPil.rkmspqNigaaRPlk

1020 1030 1040 1050 1060 1070

X.laevis RNLBMDNPLTDKKEKRAVPLHLPL..KEMSGKDTFVGODGESKSVH.DHSGS.MESHKMPLSSPFP
X.tropicalis RNLBMDNPLTEKKEKRAVPLSLP..KEMSGKDTFVGODGESKSVH.DHSGS.MMESHKMPLSSPFP
G.gallus RGLBLLDKVISEKKEKRAVPLHLPL..KETSVKKEKVPDAEGEGSKSLH.GHSGS.VMESHKMPLSAPYFP
H.sapiens RSLBLLDLVIAGKKEKRAVPLDPS..KEVSGKKEKASEQESGSKAAH.GHSGG.GMESHKMPLSPTFP
M.musculus RNLBLLDLVIAGKKEKRAVPLDPA..KEASGKKEKASEQESGKGCAY.GHSGA.AMESHKMPLSPTFP
D.ierio RNLBLLDLGPTKKEKRAVPLHCSTQRDLGKPRTPAEDSEVYKPAEPAVHINSYTESHKMPLSHHFP
consensus>50 RnlE#s#.viteKKiRAVSPHLpl..kmsgKek..eq#gEgKsah.giHsgs..sEGHkYPLssPFP

1080 1090 1100 1110 1120 1130 1140

X.laevis GMYEGSLC.G.....GLSSRIPTAYSHLQYLNKQNALSLPMLQPIALHTFMMQRQYLTNSTNSQQLYRH
X.tropicalis GMYEGSLC.G.....GLSSRIPTAYSHLQYLNKQNVLSPLMQPIALHTFMMQRQYLTNSTNSQQLYRH
G.gallus GLEYEGTLC.T.....GLNNRLPPEYSHLQYLNKQNVLSPLMQPIALHSLMVQRQFLTSPANSQQLYRH
H.sapiens GLEYEGSLCNS.....GLMSRLPAEYSHLQYLNKQNVLSPLMQPIALHSLVMQRGIFTSPANSQQLYRH
M.musculus GLEYEGSLCNS.....GLMSRLPAEYSHLQYLNKQNVLSPLMQPIALHSLVMQRGIFTSPANSQQLYRH
D.ierio GLEYEGTLEVSQVQDMCESLGSHTVPEYSHLQYLNKQNVLSPLMQPIALHSLMMQRQFLA..ANPAMVYRH
consensus>50 G\$YpGslc.....glnsrip.gYSHpLQYLNKQtlVLSPLmqPIALHslmmQRqyltspntsqqqYRH

1150 1160 1170 1180

X.laevis IASGAPVGSYGDILHNSIYPLTAINPQSPFPSSQSSSVHPSTKL.
X.tropicalis IASGAPVGSYGDILHNSIYPLTAINPQSPFPSSQSSSVHPSTKL.
G.gallus LATATPVGSYGDILHNSIYPLAAINPQAAFPPSQSSSVHPSTKL.
H.sapiens LAAATPVGSYGDILHNSIYPLAAINPQAAFPPSQSSSVHPSTKL.
M.musculus LAAATPVGSYGDILHNSIYPLAGINPQAAFPPSQSSSVHPSTKL.
D.ierio ...LQVGSYGDILHGLYPMALHPOPAPFPSPPOSSVHPSTKLS
consensus>50 Ia.atpVGSYGDILHnsiYp\$aainPQaaFpssQ\$SSVhPSTKL.

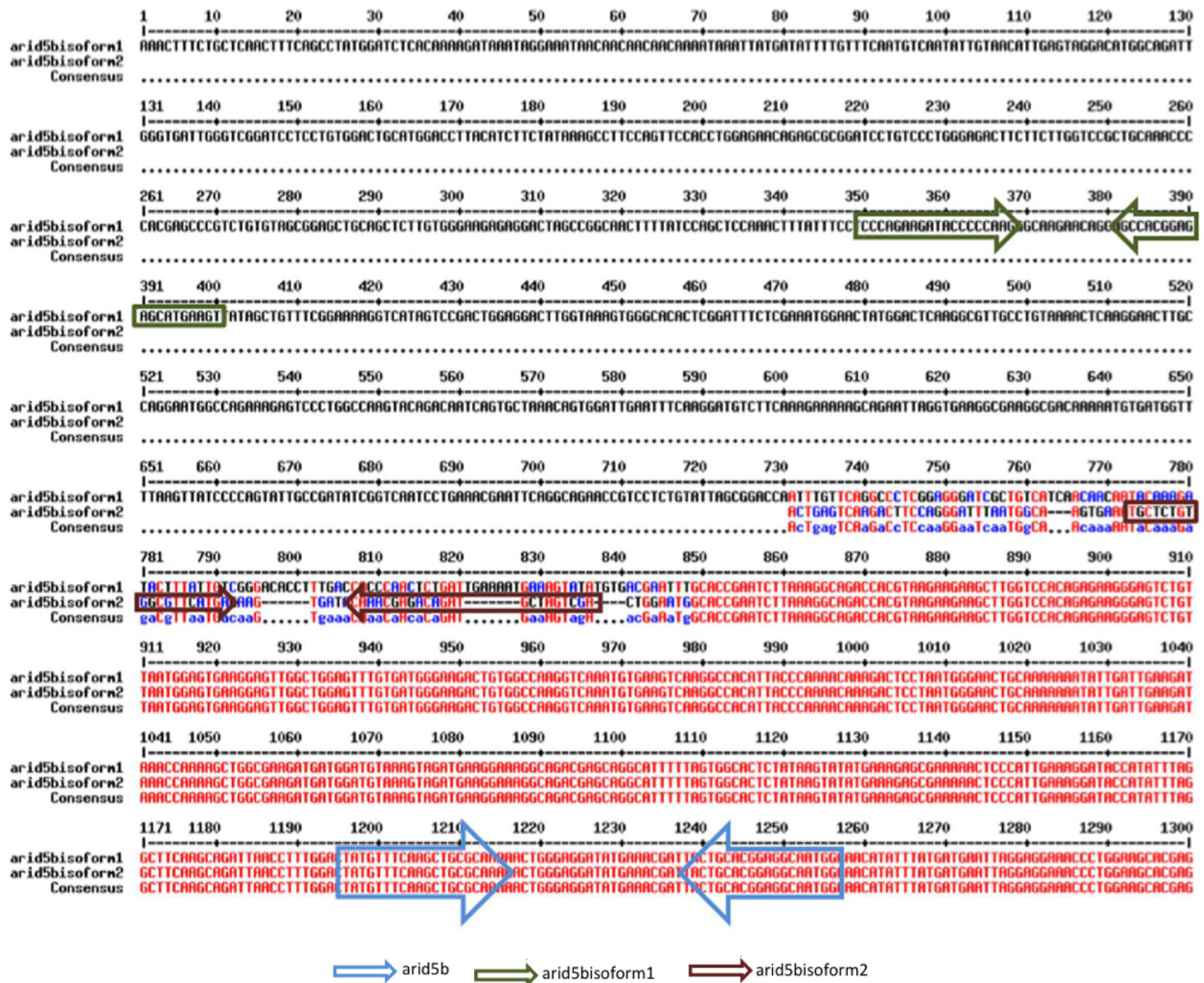
Supplementary Fig. S1. Comparison of Arid5b isoform1 amino acid sequence between vertebrates. Predicted amino acid sequence comparison of Arid5b: X. tropicalis (F6QQ73) isoform1; G. gallus (Q5ZJ69); H. sapiens (Q14865); M. musculus (Q8BM75); D. rerio (E7F888). Red boxes indicate amino acid residues conserved in all species. Phe-244 of isoform1 is replaced by a start methionine in isoform2.

```

1 80
Isoform1 MEPNSLKWVG SSCGLHGPYI FYKAFQFHLE NRARILSLGD FFLVRCKPHE PVCVAELQLL WEERTSRQLL SSSKLYFLPE
Isoform2 .....
81 160
Isoform1 DTPKGNSSH GEHEVIASE KVIVRLEDLV KWAHSDFSKW NYGLKALPVK LKELARNQKQ ESLAKYRQSV LNSGLNFKDV
Isoform2 .....
161 240
Isoform1 FKKEAELGEG EGDKNVMVLS YPQCYRYSI LKRIQAEPSV VLADQFVQAL GGIAVINNTT KILYCRDTFD HPTLIENESI
Isoform2 .....
241 320
Isoform1 CDEFAPNLKG RPRKKLGPQ RRESVNGVKE LAGVCDGKTV AKVKCEVKAT LPKTKTPNGN CKKILIEDKP KAGEDDGCKV
Isoform2 ...MPNLKG RPRKKLGPQ RRESVNGVKE LAGVCDGKTV AKVKCEVKAT LPKTKTPNGN CKKILIEDKP KAGEDDGCKV
321 400
Isoform1 DEGKADEQAF LVALYKYMKE RKTPIERIPY LGFKQINLWT MFQAAQKLGK YETITARRQW KHIYDELGGN PGSTSAATCT
Isoform2 DEGKADEQAF LVALYKYMKE RKTPIERIPY LGFKQINLWT MFQAAQKLGK YETITARRQW KHIYDELGGN PGSTSAATCT
401 480
Isoform1 RRHYERLILP YERFIGGEE KPLPSAKPRK PENGSEQEVEL KAKICGAKRI KNESQKSKKE KDPTAKGLDM TEVPPDEEDH
Isoform2 RRHYERLILP YERFIGGEE KPLPSAKPRK PENGSEQEVEL KAKICGAKRI KNESQKSKKE KDPTAKGLDM TEVPPDEEDH
481 560
Isoform1 LEADEKNMPL DYDLEETKTS VDKSESUVSE VNYPSPLEND ELEETVANKD HVTKDENSQC DPDPVDSLIIH VKEINCRQTD
Isoform2 LEADEKNMPL DYDLEETKTS VDKSESUVSE VNYPSPLEND ELEETVANKD HVTKDENSQC DPDPVDSLIIH VKEINCRQTD
561 640
Isoform1 KQLQMPNETM TTTKREIQE DYSDHLENDP EDVQLHVFFA IQAPQHDMLH EEEKLPDMPD YIANCTVKVD PLGSNDLKNP
Isoform2 KQLQMPNETM TTTKREIQE DYSDHLENDP EDVQLHVFFA IQAPQHDMLH EEEKLPDMPD YIANCTVKVD PLGSNDLKNP
641 720
Isoform1 LDSNLLQNAL KQNPKVYFVQ TLDMLSDEKD TSASMNDSS FSYTPLLYSR GNPGIMSPLA KKKLLSQVSG ASQPGNLPYG
Isoform2 LDSNLLQNAL KQNPKVYFVQ TLDMLSDEKD TSASMNDSS FSYTPLLYSR GNPGIMSPLA KKKLLSQVSG ASQPGNLPYG
721 800
Isoform1 SPPPLISKKK LSSKGEVSPS LLQTHSSNS ESAAINRPSV IQHVQSFQKQ SPEEKKTVDN HYKNMFMGKV DSYCCDFAKH
Isoform2 SPPPLISKKK LSSKGEVSPS LLQTHSSNS ESAAINRPSV IQHVQSFQKQ SPEEKKTVDN HYKNMFMGKV DSYCCDFAKH
801 880
Isoform1 HQSVLADSYA LKSCVQECKE KMAEKRAASN SNVPSFVAEF YSSPHLHRLY RQAEHHLHNE NSAKFHSREM FRDLENVSSH
Isoform2 HQSVLADSYA LKSCVQECKE KMAEKRAASN SNVPSFVAEF YSSPHLHRLY RQAEHHLHNE NSAKFHSREM FRDLENVSSH
881 960
Isoform1 KHYYHASLHQ HDKQNLHDDV DDQPTDLSLP KSLHLKSTKI PGSSICHQPV QDQSKSHNPF QTPNSKTGLG DCPNPKACRVS
Isoform2 KHYYHASLHQ HDKQNLHDDV DDQPTDLSLP KSLHLKSTKI PGSSICHQPV QDQSKSHNPF QTPNSKTGLG DCPNPKACRVS
961 1040
Isoform1 PMTMPISKRH MDSIQRPSKT VKPDTLRKVE GLVHPFSIGK TNTHNFGAPR SLKRNLDMND NPLTDKKIRA VSPLHLPKEM
Isoform2 PMTMPISKRH MDSIQRPSKT VKPDTLRKVE GLVHPFSIGK TNTHNFGAPR SLKRNLDMND NPLTDKKIRA VSPLHLPKEM
1041 1120
Isoform1 SGKDTFVGQD GESSKSVHDI HSGSMIESHK YPLSSPFFPG MYPGSLCGGL SSRIPYAYSH PLQYLKNQTA LSPMLQPLAL
Isoform2 SGKDTFVGQD GESSKSVHDI HSGSMIESHK YPLSSPFFPG MYPGSLCGGL SSRIPYAYSH PLQYLKNQTA LSPMLQPLAL
1121 1187
Isoform1 HTFMMQRQYL TNSTNSQQLY RQIASHAPVG SSYGDLHSS IYPLTAINFQ SPFPSSQMSS VYFSTKL
Isoform2 HTFMMQRQYL TNSTNSQQLY RQIASHAPVG SSYGDLHSS IYPLTAINFQ SPFPSSQMSS VYFSTKL

```

Supp. Fig. S2. Amino acid comparison of the two Arid5b isoforms. The ARID domain is highlighted in yellow. Regions specific to isoform1 or isoform2 are indicated in red letters. The sequence in blue corresponds to the predicted C-terminal sequence of isoform2 observed in 3'RACE PCR experiment but that we could not amplify by end-to-end PCR.



Supp. Fig. S3. Nucleotide position of PCR primers used for RT-qPCR. Nucleotide sequences of the N-terminal region of arid5b isoform1 and isoform2 are aligned. Arrows indicate the primer sequences that were used to amplify both isoforms (in blue), isoform1 (green) and isoform2 (brown). Conserved nucleotides are in red.

TABLE S1

**EXON-INTRON ORGANIZATION OF THE *ARID5B* GENE
IN *XENOPUS LAEVIS* ON SCAFFOLD 9729**

Exon n°	Exon size (bp)	Sequence at exon-intron junction		Intron size (bp)	Amino acid interrupted
		5' splice donor	3' splice acceptor		
1	>27	ACTCAAG <u>gt</u> tattattagctg	tcttgc <u>ag</u> TGGGTCGG	328	Trp-8
2*	>68	?	atttc <u>ag</u> CATGAAGT	>39437	Glu-92?
3	226	GAATTAG <u>gt</u> aaccatgggttc	aaccac <u>ag</u> TGAAGGCG	40399	His-93
4	231	AATTTGG <u>gt</u> taagtccaatc	ttttac <u>ag</u> CACCGAAT	66835	Gly-168
5	113	GGCCAAG <u>gt</u> aaatcctttt	tctttc <u>ag</u> GTCAAATG	1316	Lys-282
6	220	AAGCAGAG <u>gt</u> aagtagacact	tctaac <u>ag</u> TTAACCTT	4458	Ile-356
7	53	TGAAACG <u>gt</u> taagtcttta	tataac <u>ag</u> ATTACTGC	3545	Ile-374
8	98	ACGAAAG <u>gt</u> taggataatcta	ccccac <u>ag</u> ATTAATCC	33496	Arg-406
9	199	GACTGAG <u>gt</u> aaattggggca	ttgtt <u>ag</u> GTTCCACC	5264	Val473
10	>2712				
4b	>99	GGAATGG <u>gt</u> aacogcttcag	tctttc <u>ag</u> GTCAAATG	6493	Ala-2 (short isoform)

Exon sequences are indicated by uppercase letters and intron sequences by lowercase letters. Splice donor and acceptor sites are underlined. Exon and intron size are reported as base pairs.

Scaffold 9729. * sequence nt89-275 mRNA ORFmissing in genomic sequence.

TABLE S2

**EXON-INTRON ORGANIZATION OF THE *ARID5B* GENE
IN *XENOPUS LAEVIS* ON SCAFFOLD 48311**

Exon n°	Exon size (bp)	Sequence at exon-intron junction		Intron size (bp)	Amino acid interrupted
		5' splice donor	3' splice acceptor		
1*	>27	ACTCAAG <u>gt</u> tattattagctg	?	?	Trp-8
2	236	CGGAGAG <u>gt</u> gacatccctac	ttttc <u>ag</u> CATGAAGT	53173	His-93
3	226	GAACTAG <u>gt</u> aaccatgattc	aaccac <u>ag</u> GTGAAGGC	53008	Gly-168
4	231	GAATTTG <u>gt</u> taagttccaatt	ttttc <u>ag</u> CACCAAAT	77336	Ala-245
5	113	TGCCAAG <u>gt</u> aaatcggtttt	tgtttt <u>ag</u> GTCAAATG	1264	Val-283
6	220	AAGCAGAG <u>gt</u> aagtagacatt	tctaac <u>ag</u> TTAACCTT	4033	Ile-356
7	53	TGAAACG <u>gt</u> taagtctttt	tgtaac <u>ag</u> ATTACTGC	4690	Ile-374
8	98	ACGAAAG <u>gt</u> taggataatctg	ccccac <u>ag</u> ATTAATCC	33865	Arg-406
9	199	GGCTGAG <u>gt</u> tattggggca	ttgtt <u>ag</u> GTTCCACC	4311	Val-473
10	>2705				
4b	>96	GGAATGG <u>gt</u> aacogctttat	ttttc <u>ag</u> CACCAAAT	5945	Ala-2 (short isoform)

Exon sequences are indicated by uppercase letters and intron sequences by lowercase letters. Splice donor and acceptor sites are underlined. Exon and intron size are reported as base pairs.

Scaffold 48311 * sequence nt21-39 mRNA ORFmissing in genomic sequence.